

What is claim d is:

1. A document processing device for processing a set having, as elements, a plurality of documents including character information, comprising:

a theme information generating unit for extracting the commonality of the character information of the respective documents in the set to generate theme information which is a document expressing common semantic contents being common in the entire set.

2. A document processing device according to claim 1, wherein the theme information generating unit executes a predetermined generating procedure on the basis of the plurality of documents in the set to generate the theme information as a new document.

3. A document processing device according to claim 2, wherein the theme information generating unit extracts an expression being common in the plurality of documents in the set and generates the theme information on the basis of the expression.

4. A document processing device according to claim 1, wherein the theme information generating unit selects documents having a commonality in character information in advance to constitute the set, selects one document from the plurality of documents in the set depending on a predetermined selecting procedure, and sets the selected document as the theme information to generate the theme information.

5. A document processing device according to claim 4, wherein in the selecting procedure, frequent expressions which appear times the number of which is not less than a predetermined threshold value in the documents in the set are detected, and a document which includes the frequent expressions the number of which is largest is selected.

6. A document processing device according to claim 1, comprising

a difference character information generating unit for generating difference character information serving as character information serving as differences between the theme information and the documents in the set.

7. A document processing device according to claim 6, comprising:

an information presenting unit for displaying a document obtained by the theme information generating unit; and

a text set accumulating unit for accumulating the documents processed into documents of a predetermined display form suitable for display on the information presenting unit, wherein

the information presenting unit displays that the difference character information is differences between the theme information and the documents in the set.

8. A document processing device according to claim 7, wherein the information presenting unit displays the theme information and difference character information between the theme information and the documents on a screen depending on an operation by a user.

9. A document processing device according to claim 8, wherein the information presenting unit displays source information representing an on-line or off-line source of a document including the difference character information on a screen depending on an operation by the user.

10. A document processing device according to claim 9, wherein when the user performs a predetermined operation to the source information displayed on the screen, the information presenting unit executes predetermined identification display for identifying difference character information between a document corresponding to the source

information and the theme information to the theme information displayed on the screen.

11. A document processing method for processing a set having, as elements, a plurality of documents including character information, wherein

a theme information generating unit extracts the commonality of the character information of the respective documents in the set to generate theme information which is a document expressing common semantic contents being common in the entire set.

12. A document processing method according to claim 11, wherein

the theme information generating unit executes a predetermined generating procedure on the basis of the plurality of documents in the set to generate the theme information as a new document.

13. A document processing method according to claim 12, wherein

the theme information generating unit extracts an expression being common in the plurality of documents in the set and generates the theme information on the basis of the expression.

14. A document processing method according to claim 11, wherein

the theme information generating unit selects documents having a commonality in character information in advance to constitute the set, selects one document from the plurality of documents in the set depending on a predetermined selecting procedure, and sets the selected document as the theme information to generate the theme information.

15. A document processing method according to claim 14, wherein

in the selecting procedure, frequent expressions which appear

times the number of which is not less than a predetermined threshold value in the documents in the set are detected, and a document which includes the frequent expressions the number of which is largest is selected.

16. A document processing method according to claim 11, wherein

difference character information serving as character information serving as differences between the theme information and the documents in the set is generated.

17. A document processing method according to claim 16, comprising:

an information presenting unit for displaying a document obtained by the theme information generating unit displays that the difference character information is differences between the theme information and the documents in the set.

18. A document processing method according to claim 17, wherein

the information presenting unit displays the theme information and difference character information between the theme information and the documents on a screen depending on an operation by a user.

19. A document processing method according to claim 18, wherein

the information presenting unit displays source information representing an on-line or off-line source of a document including the difference character information on a screen depending on an operation by the user.

20. A document processing method according to claim 19, wherein

when the user performs a predetermined operation to the source information displayed on the screen, the information presenting unit

executes predetermined identification display for identifying difference character information between a document corresponding to the source information and the theme information to the theme information displayed on the screen.